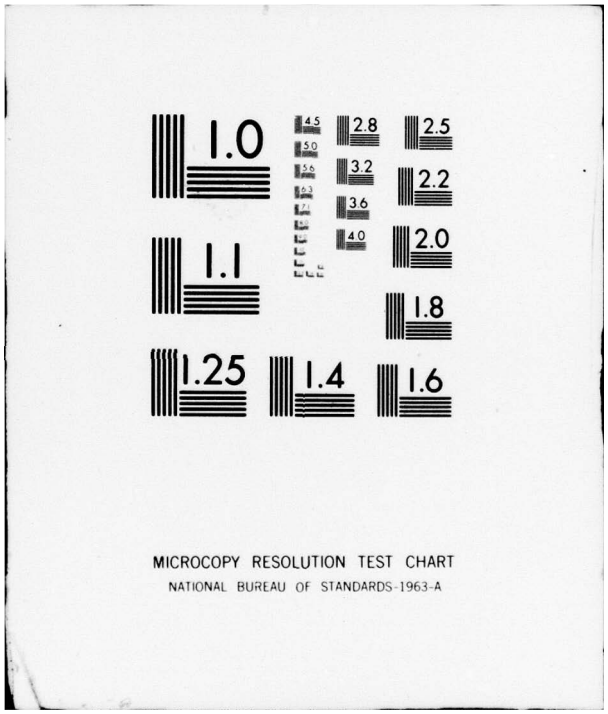
AD-A081 188 CHICAGO UNIV ILL DEPT OF GEOPHYSICAL SCIENCES F/G 12/1
RECONSTRUCTIVE TOMOGRAPHY: AN INVERSE PROBLEM.(U)
DEC 79 C J LEVINE N00014-76-C-0034
UNCLASSIFIED NL

| OF |
AD-
A081188

END
DATE FILMED
3-80
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

THE UNIVERSITY OF CHICAGO

RECONSTRUCTIVE TOMOGRAPHY: AN INVERSE PROBLEM.

(9) Master's thesis,

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF GEOPHYSICAL SCIENCES

(12) 49

(15) N0014-76-C-0034

BY

(10) CATHERINE J. LeVINE

CHICAGO, ILLINOIS

(11) DECEMBER, 1979

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author would like to thank Professor Victor Barcilon for suggesting this paper, and for encouragement and financial support during its preparation.

The author also thanks Dipl. Geol. Juliane Fenner for translating the paper by J. Radon.

The author also thanks Marilyn Bowie for typing the manuscript.

The author would like to express special appreciation to Stephen Daly, David Thorn and Robert Haskins for their continued encouragement and moral support.

This research was supported by the Office of Naval Research under contract N00014-76-C-0034. Reproduction in whole or in part is permitted for any purpose of the United States Government.

LIST OF ILLUSTRATIONS

INTRODUCTION

The process of recovering the three-dimensional structure of an object by reconstructing successive cross sections orthogonal to a common axis is known as reconstructive tomography. Each cross section of the object is essentially a function f defined on a two-dimensional domain C. This unknown function may be specified using as data integrals of f along straight lines traversing C. Data of this sort represent cumulative measures of f along certain paths rather than actual values of f at specific points. Suppose f vanishes outside the unit disk D: $x^2 + y^2 \leq 1$; since any finite domain may be enclosed by a disk this is not a restriction on C. Let $L(t,\theta)$: $x \cos\theta + y \sin\theta = t$ denote the straight line which makes an angle θ with the positive y axis and whose perpendicular distance to the origin is t, as shown in Fig. 1. The integral of f along $L(t, \theta)$ is

$$p(t, \hat{\xi}) = \int_{L(t,\theta)} f(\underset{\sim}{r})\, ds \qquad (1)$$

where $\underset{\sim}{r} = (x, y)$ and $\hat{\xi}$ is the unit vector $\hat{\xi} = (\cos\theta, \sin\theta)$. The original coordinates x,y are related to the rotated coordinates t,s by

$$\begin{aligned} x &= t\cos\theta - s\sin\theta \\ y &= t\sin\theta + s\cos\theta. \end{aligned} \qquad (2)$$

Equivalently, the integral of f along $L(t,\theta)$ may be expressed as

$$p(t,\hat{\xi}) = \int_D f(\underset{\sim}{r})\ \delta(\underset{\sim}{r}\cdot\hat{\xi} - t)\ d^2r \qquad (3)$$

where $\underset{\sim}{r}\cdot\hat{\xi} = x\cos\theta + y\sin\theta$ and $\delta(r)$ is the Dirac delta function. The set of line integrals for various values of t, $|t| \leq 1$, at a given angle θ is known as a scan of f. When, for a given θ, the integral values are specified for all t, $|t| \leq 1$, the scan will be called a projection of f onto the line orthogonal to the direction θ. In much of the following discussion the line integral data will be assumed to be projections of f, that is, continuous in t. A complete data set includes projections for all angles $\theta\in[0,2\pi)$. This problem has applications in a number of fields.

The reconstruction of the human body has received much attention in the last decade. Computerized Axial Tomography (CAT) transmission scanners use measurements of transmitted radiation such as x-rays, gamma rays or protons to determine the corresponding attenuation coefficient and, by inference, the density of a cross section (Cormack 1963, 1964; Cho *et al.* 1974). The data represent the total attenuation of a beam of radiation due to its passage through an object. Emission scanners reconstruct the isotope distribution of a cross section from measurements of radionuclide emissions (Budinger and Gullberg 1974b, Kuhl and Edwards 1968, Kuhl *et al.* 1973). Most scanners operate in the parallel mode in which the line integral data are obtained by taking measurements at successive positions of the detector as it is translated parallel to the t axis as shown in Fig. 1. In the diverging beam or fan beam scanners the object is scanned by radiation diverging from a point source S as shown in Fig. 2. With either configuration the source-detector apparatus is rotated about the object or, equivalently, the object is rotated within the apparatus

to produce a set of line integral scans. Since the density or the isotope concentration of abnormal tissue differs from that of healthy tissue, such reconstructions are a useful diagnostic tool. They are especially important in the detection of brain tumors, a task for which conventional x-rays are of little use.

The three-dimensional structure of biomolecules may be determined from electron micrographs. Rotation of the specimen within the microscope results in a set of two-dimensional maps of the total attenuation of the electron beam as it passes through the particle in different orientations. The three-dimensional optical density of the specimen may be reconstructed directly from this set of two-dimensional projections (Hart 1968; DeRosier and Klug 1968; Crowther *et al.* 1970). In an alternative method (Gordon *et al.* 1970; Vainshtein 1970, 1971), each micrograph is regarded as a set of parallel lines; each line is a measure of the attenuation of the beam due to passage through a particular cross section of the particle. Each cross section is reconstructed from the set of these one-dimensional projections, one from each micrograph, representing different orientations of the particle. The three-dimensional density structure is then recovered by stacking the cross sections.

In radio astronomy the two-dimensional brightness distribution of radio sources or the Sun or Moon are inferred from measurements of the intensity of radiation received by radio telescopes (Bracewell and Roberts 1954, Bracewell and Riddle 1967). The telescope collects radiation along a narrow strip which may be idealized as a line. The amount of radiation received for a given orientation of the antenna is a relative measure of the total radiation emitted along a similarly-positioned line across the

surface of the source. Adjustment of the telescope or the rotation of the Earth results in a scan of lines at the same angle. A set of scans is obtained by changing the orientation of the antenna. Records of a lunar occultation of a star will produce similar scans and the star's brightness distribution may be reconstructed from data for occultations occurring from different directions.

The use of travel time data for waves propagating through a medium allows properties of the medium to be inferred. Such data represent the integrals of the slowness, the reciprocal of the velocity, along the paths of the waves. In seismology travel times traditionally have been used to determine the vertical seismic velocity, and hence density, structure of the earth. Recently Aki *et al.* (1977) developed a method for finding the three-dimensional slowness structure. They assumed an initially layered earth model and divided each layer into blocks within which a perturbation to the initial slowness model is assumed to be constant. The travel time residual, the difference between the observed travel time and that calculated for the initial model, for each seismic ray received is then the sum of the slowness parameters for each block through which the ray has passed. With this method the problem is solved in three dimensions. Munk and Wunsch (1979) treated the problem of determining mesoscale fluctuations in the density of the ocean from travel times for acoustic waves in a similar manner. They used the block structure of Aki *et al.* to reconstruct the sound speed perturbation within either horizontal or vertical cross sections and stacked the cross sections to obtain the three dimensional structure. Their data, based on an array of several acoustic sources and receivers, was composed of all the resolvable waves arriving at each receiver along all possible ray paths.

Some of these applications involve the reconstruction of a three-dimensional function from a set of two-dimensional projections or the reconstruction of two-dimensional functions from line integral data rather than the reconstruction of a three-dimensional function by stacking a sequence of cross sections obtained from line integral data. These applications are not considered tomography in the strict sense but techniques devised for the solution of these problems have been applied to the tomography problem as well.

Each of the problems described above may be characterized by the formulation in Eq. (3). This equation, restated in canonical form, is

$$p(t,\hat{\xi}) = Lf = \int_D G(\underset{\sim}{r},\hat{\xi};t)\ f(\underset{\sim}{r})\ d^2r \tag{4}$$

where $G(\underset{\sim}{r},\hat{\xi};t) = \delta(\underset{\sim}{r}\cdot\hat{\xi} - t)$ and t is a parameter. Since the operator L is linear the problem of solving this equation for f may be approached using well-developed techniques in inverse theory. The prediction of the data $p(t,\hat{\xi})$ when $f(\underset{\sim}{r})$ is known is the direct or forward problem. In general the direct problem involves the formulation of a model relating some observable property of a function to the function itself. The appropriateness of the model is evaluated by comparison of calculated data with observed data. The tomography problem involving transmitted radiation, for example, is based on the observed exponential attenuation of radiation due to passage through matter. The intensity $I(t,\hat{\xi})$ of a transmitted beam of radiation is known to be related to the linear attenuation coefficient $\mu(\underset{\sim}{r})$ of the medium through which it has passed by

$$I(t,\hat{\xi}) = I_o \exp\left[- \int_{L(t,\theta)} \mu(\underset{\sim}{r})\ ds\right] \tag{5}$$

in which I_o is the intensity of the incident ratiation. The data

$$p(t,\hat{\xi}) = - \ln[I(t,\hat{\xi})/I_o] \tag{6}$$

are thus related to $\mu(\underset{\sim}{r})$ by Eq. (1). A similar formulation of a model which relates a function f to a measurable property of f is more difficult because the unknown is a function rather than a collection of numbers. For this reason solution of the inverse problem is generally attempted by inverting the operator L, resulting in

$$f(\underset{\sim}{r})=L^{-1}[p(t,\hat{\xi})] \tag{7}$$

where L^{-1} is the operator inverse to L. In most problems this is not an easy task either. In addition to the construction of a solution which satisfies Eq. (4) for a specified data set, inverse theory addresses such questions as the existence, uniqueness and stability of possible solutions.

The purpose of this paper is to review the literature on reconstructive tomography within the framework of inverse theory to be described below. The concepts of existence, uniqueness, construction and stability of solutions to Eq. (4) will be explored and discussed in terms of their application to tomography.

Proof of the existence of a solution to an inverse problem ensures that there is a solution which is consistent with the problem as posed; it does not necessarily imply that a solution has been found. Often the assumptions incorporated into the formulation of a problem impose constraints on the data. If these consistency conditions are violated there is no solution which satisfies both the data and the original assumptions. In the tomography problem, for example, if the solution is assumed to be radially symmetric its scans should be the same for all angles. If the

measured scans are not in fact independent of θ, the assumption of radial symmetry is invalid. The consistency of a data set with such constraints does not, however, guarantee the existence of a solution. In many problems a solution which is consistent with the data is assumed to exist.

If a solution to an inverse problem does exist, it may not be the only solution which satisfies the given data. Possible sources of such non-uniqueness include contamination of the data by errors or, in the case of perfect data, the finiteness of the data set. The non-uniqueness associated with a partial data set may be motivated in terms of the tomography problem as follows: There are an infinite number of functions which satisfy a data set composed of only one scan. A given line integral value $p(t, \hat{\xi})$ might be produced, for instance, by a point source at any of the points along $L(t, \theta)$; there is no way to distinguish between these possible solutions. If a second scan is introduced the number of solutions which produce both scans, while still infinite, is reduced. Only when scans are given for all angles $\theta\varepsilon[0,2\pi)$ is a unique solution possible. For some inverse problems even a complete, error-free data set is insufficient to resolve the non-uniqueness. See Backus (1970) for an example of such inherent non-uniqueness. Even though a solution is completely undetermined by a finite data set, some information may be extracted. Since obtaining a unique solution is generally not possible a reasonable alternative is to attempt to characterize the class of solutions which satisfy the given data. Bounds on some property, such as the minimum value, of the solution may be established by selecting from this class that solution which extremizes the property of interest. Such extremal solutions are unique (Parker 1972, 1975; Sabatier 1977a, 1977b).

An inverse problem is considered well-posed or stable if for each set of data a solution exists, is unique and depends continuously on the data. Continuous dependence on the data implies loosely that a small change in the data will produce a small change in the solution. This means that an error in the data will not cause the solution to deviate substantially from the solution which would be obtained if the data were perfect. Linear inverse problems on a finite domain described by an integral equation of the form (4) are generally unstable if the kernel G is continuous.

There is often more than one method of constructing a solution to an inverse problem. This is especially true for the tomography problem. Construction procedures are generally judged on their accuracy and the computation time they require.

This paper will focus on reconstructive tomography as an inverse problem. There are other reviews which emphasize different aspects of this subject. Brooks and DiChiro (1976) reviewed several reconstruction methods and discussed experimental and diagnostic limitations of tomography. Gordon and Herman (1974) presented a comparative survey of different reconstruction techniques. Reviews by Smith *et al.* (1977) and Shepp and Kruskal (1978) described contributions to the mathematical theory of tomography.

CHAPTER I

EXISTENCE AND UNIQUENESS OF SOLUTIONS

The operator L defined in Eq. (4) is known as the Radon transform after Johann Radon who originally solved the reconstruction problem (Radon 1917). Radon formed an average $\bar{p}(r;q)$ of the line integral values at a distance q from $\underset{\sim}{r} = (x,y)$, weighted it by q^{-2}, and summed over all possible values of q to obtain the value of the source function at $\underset{\sim}{r}$. His inversion formula is

$$f(\underset{\sim}{r}) = -\frac{1}{\pi}\int_0^\infty \frac{d\bar{p}(\underset{\sim}{r};q)}{q} \tag{8}$$

$$= \frac{1}{\pi}\lim_{\varepsilon\to 0}\left\{\frac{\bar{p}(\underset{\sim}{r};\varepsilon)}{\varepsilon} - \int_\varepsilon^\infty \frac{\bar{p}(\underset{\sim}{r};q)}{q^2}\,dq\right\} \tag{9}$$

where

$$\bar{p}(\underset{\sim}{r};q) = \frac{1}{2\pi}\int_0^{2\pi} p(\underset{\sim}{r}\cdot\hat{\xi} + q,\hat{\xi})\,d\theta. \tag{10}$$

Weighting the average by q^{-2} emphasizes the contribution of line integrals near $\underset{\sim}{r}$ and lessens that of those at a distance. Radon assumed that the line integral scans $p(t,\hat{\xi})$ are continuous in t, are specified for all angles $\theta\in[0,2\pi)$, have continuous derivatives in t and θ, and that both the scans and their derivatives satisfy certain conditions on their asymptotic behavior. Under these assumptions he proved that the function $f(\underset{\sim}{r})$ obtained by this inversion formula exists, is real and continuous, and is

unique. Thus in the tomography problem the non-uniqueness associated with a partial data set is resolved when a complete data set is given.

Radon also generalized the inversion to n dimensions using as data integrals over (n-1)-dimensional hyperplanes. If n is even the reconstruction of f at a point $\underset{\sim}{r}$ involves all the line integral data whereas if n is odd only those integrals along lines $L(t,\theta)$ in a small neighborhood of $\underset{\sim}{r}$ are required (Gel'fand *et al.* 1966). The reconstructions discussed in this paper will be two- and three-dimensional.

In order for a set of functions $\{g(t,\hat{\xi}): \theta\in[0,2\pi)\}$ to represent scans of a solution f the functions $g(t,\hat{\xi})$ must satisfy certain consistency conditions imposed by the form of the operator L. One such condition is that $g(t,\hat{\xi}) = g(-t,-\hat{\xi})$, that is, that the integral of f along any line $L(t,\theta)$ must be independent of the direction of integration. Secondly, the functions $g(t,\hat{\xi})$ should belong to the same space, with a different domain of definition, as the function f to be reconstructed. For example, a function $g(t,\hat{\xi})$ representing a scan of a function which vanishes outside the unit disk must vanish for $|t| > 1$. A final condition is that $\int_{-\infty}^{\infty} g(t,\hat{\xi})t^k\,dt$ be a polynomial in $\hat{\xi}$ of degree less than or equal to any positive integer k. This condition is a consequence of the equivalence

$$\int_{-\infty}^{\infty} (Lf)(t,\hat{\xi})\ t^k\,dt = \int_D f(\underset{\sim}{r})\ (\underset{\sim}{r}\cdot\hat{\xi})^k\,d^2r \qquad (11)$$

in which the right side is a polynomial in $\hat{\xi}$ of degree less than or equal to k. These conditions are both necessary and sufficient for the existence of a function f such that $g = Lf$ (Ludwig 1966). Other consistency conditions arise in connection with certain reconstruction methods and will be mentioned as these methods are described.

The uniqueness result obtained by Radon is in general valid only when scans at all angles $\theta\varepsilon[0, 2\pi)$ are available. One exception occurs when the unknown function is known to be radially symmetric. In this case one scan which is continuous in t is equivalent to the complete set of scans and $f(\underset{\sim}{r})$ is determined uniquely (Bracewell 1956). A second exception provides a measure of the resolution which may be achieved in a reconstruction based on a finite number of scans. Suppose a function $f(\underset{\sim}{r})$ is constant on square grid elements with side length b and vanishes for $|x| > a$, $|y| > a$, where a is a multiple of b. Then $f(\underset{\sim}{r})$ may be reconstructed uniquely from a set of scans at angles θ_j, $j=1, \ldots, 2a/b$, such that $0 < \theta_j < \tan^{-1} 1/(j-1) \leq \pi/2$ (Frieder and Herman 1971). This restriction on the spacing of the angles, which rules out a unique reconstruction if the angles are equally spaced, is actually an advantage in electron microscopy where the range of available angles is generally very narrow. The scans may be sampled discretely in t if the sampling interval is less than $\sqrt{2}\,b$. Any feature smaller than b which appears in a reconstruction must be an artifact of the reconstruction procedure.

When line integral scans at a finite number of angles $\theta_j, j=1, \ldots, N$, are given the reconstruction problem may be stated as

$$p(t,\hat{\xi}_j) = L_j f = \int_D f(\underset{\sim}{r})\ \delta(\underset{\sim}{r}\cdot\hat{\xi}_j - t)\ d^2r \qquad j = 1, \ldots, N \tag{12}$$

where $\hat{\xi}_j = (\cos\theta_j, \sin\theta_j)$. In the absence of *a priori* knowledge concerning the symmetry properties of the unknown f, the data set $\{p(t,\hat{\xi}_j)\}_{j=1}^{N}$ is not sufficient to specify f uniquely (Solomon 1976). There will always be functions which integrate to zero in the directions θ_j. Solomon proved that any function in the null space of L_j is orthogonal to all functions

which are constant along the lines $L(t,\theta_j)$, $|t| \leq 1$, and gave a formula for the angles between the null spaces in terms of the angles θ_j. Bracewell and Roberts (1954) also showed that there are "invisible distributions" which cannot be resolved by a finite number of projections. They defined a principal solution to be that solution which satisfies the given data and for which the null spaces are assumed to be zero.

Use of a sampling theorem for Fourier transforms allows this principal solution to be determined uniquely from a finite number of scans. A function may be completely recovered by a discrete sampling of its two-dimensional Fourier transform if the transform vanishes outside a circle of radius R_o, known as the cut-off frequency. If the source function $f(\underline{r})$ is confined within a circle of radius r_o, the sampling interval should be less than r_o^{-1} (Bracewell and Riddle 1967). Knowledge of the scans of f in the directions θ_j, j=1, ..., N, is equivalent to knowledge of the two-dimensional Fourier transform of f along radial lines at the same angles, as will be detailed below. If the number of angles is chosen so that the spacing of the radial lines near $R=R_o$ is within the sampling interval r_o^{-1} then a principal solution for f is completely determined, up to a component corresponding to spatial frequencies greater than R_o. This idea, developed in radio astronomy, is based on the properties of radio telescopes which cannot resolve frequencies beyond some cut-off frequency R_o. Although the Fourier transform of an actual distribution may not vanish for $R > R_o$, there is no loss of information involved in making this assumption since such frequencies cannot be measured. The principal solution represents all the information which may be obtained in this manner from a finite set of scans. If the amplitude of the

transform is small near the cut-off frequency, implying a nearly continuous transition to zero, the principal solution will be a good approximation to the true solution (Bracewell and Roberts 1954). Since one of the necessary conditions for the existence of a solution to the tomography problem is that $p(t,\hat{\xi}) = 0$ for $|t| > 1$, the two-dimensional Fourier transform of f, determined from the one-dimensional transform of $p(t,\hat{\xi})$, will, in most instances, be small near R=1 and for R > 1 (Bracewell 1965, Ch. 18). As a result the principal solution obtained from a finite set of scans will be a reasonable approximation to the actual solution. If the amplitude of the transform is not small near R=1 the principal solution will not be a good approximation to the true solution; the discontinuity in the transform produces oscillations in the solution.

Logan (1975) expressed this result more quantitatively. Let $\varepsilon(\underset{\sim}{r}) = f(\underset{\sim}{r}) - f_N(\underset{\sim}{r})$ where $f(\underset{\sim}{r})$ is the true solution and $f_N(\underset{\sim}{r})$ an approximation based on a set of N scans. Let $E(\underset{\sim}{R})$ and $F(\underset{\sim}{R})$ denote the Fourier transforms of $\varepsilon(\underset{\sim}{r})$ and $f(\underset{\sim}{r})$, respectively. If $F(\underset{\sim}{R})$ is concentrated within a circle of radius R_c, define $\lambda_N(R_c)$ as the fraction of the total energy of $E(\underset{\sim}{R})$ which is contained within the same circle. This cut-off frequency R_c is the same as that mentioned above if $F(\underset{\sim}{R})$ actually vanishes for $R > R_o$; otherwise $R_c > R_o$. If little of the energy of $E(\underset{\sim}{R})$ is contained within R_c, that is, if $\lambda_N(R_c)$ is small, then ε is small and f is approximately reconstructed by N scans. If $\lambda_N(R_c)$ is near one then ε has a large component on the null space of the operator L and f_N is a poor approximation to f. Logan showed that

$$\lim_{N\to\infty} \lambda_N(N + \alpha N^p) = 1 \qquad \alpha > 0,\ p > 1/2$$

$$\lim_{N\to\infty} \lambda_N(N - \alpha N^p) = 0 \qquad \alpha > 0,\ p > 1/3. \tag{13}$$

Thus a set of N scans can resolve a function $f(\underset{\sim}{r})$ whose Fourier transform is contained within a circle of radius R_c which is somewhat less than N. Conversely, if the Fourier transform of f is known to be contained within a circle of radius R_c, then N, somewhat greater than R_c, projections are needed to obtain a good approximation to f.

A finite data set, although generally insufficient to specify a unique solution to an inverse problem, does limit the class of possible solutions. One means of characterizing this reduced class of viable solutions is through extremal solutions. Parker (1972) presented a variational approach for finding bounds on various properties of possible model solutions, such as an integral of the model, its norm or its maximum value. In his terminology (Parker 1975), the ideal solution is that solution which extremizes some given property of the model. For the reconstruction problem Logan and Shepp (1975) proved the uniqueness of the ideal solution having the minimum norm and showed it to be of the form

$$f(\underset{\sim}{r}) = \sum_{j=1}^{N} \rho_j(\underset{\sim}{r}\cdot\hat{\xi}_j). \tag{14}$$

The ρ_j's, known as ridge functions, are constant along the lines $L(t,\theta_j)$ and vary only in t. They gave an explicit form for $f(\underset{\sim}{r})$ when the angles θ_j are equally spaced. Marr (1974) developed a least squares polynomial solution to the diverging beam tomography problem. He reconstructed a function on the unit disk using as data line integrals along the M(M-1)/2 chords which connect each pair of M equally spaced points on the unit circle. He proved that within the space of polynomials the solution which best fits the data in a least squares sense is a unique polynomial in x and y of total degree $d \leq M-2$.

CHAPTER II

RECONSTRUCTION METHODS

Radon's inversion formula is impractical for implementation because it is difficult to choose a discretization so that there are enough integral values for lines at a distance q from $\underline{r} = (x,y)$ to give a good average (Shepp and Kruskal 1978). As a result many methods for the approximate reconstruction of functions from their line integrals have been devised. In most methods the line integral data, given for a finite number of angles θ_j, j=1, ..., N, will be assumed to be continuous in t.

Summation methods

The summation methods involve, as the name implies, a superposition of the line integral data. Hart (1968) used a photographic superposition in electron microscopy. The general shape of biological particles is inferred from the bright region where the micrographs overlap. A second type of summation is known as back projection (Vainshtein 1970). In this method, also originally developed in electron microscopy, the line integral scans are regarded as the projections of the object to be reconstructed. The projections at different angles are equidistant from the center of the object and may be thought of as tangents to a circle K concentric to and enclosing the domain D. The point of tangency of each projection is the point t=0. For each line $L(t,\theta_1)$ in the scan at the angle θ_1 the corresponding value of $p(t,\hat{\xi}_1)$ is assigned to each

point along that line within the circle K. This process is called back projecting. The values of $p(t,\hat{\xi}_2)$ are assigned in a similar manner; at those points along $L(t,\theta_2)$ to which a value was assigned during the back projection of $p(t,\hat{\xi}_1)$ the appropriate value of $p(t,\hat{\xi}_2)$ is added to the previously assigned value. This procedure is continued until each scan has been back projected. The value of the reconstructed function at a point $\underline{r}$ is then the sum of all the line integral values $p(t,\hat{\xi}_j)$ corresponding to t's for which $L(t,\theta_j)$ intersects $\underline{r}$. Kuhl and Edwards (1968) adapted this method slightly by averaging the line integral values assigned to each point. A major drawback of both the photographic superposition and the back projection methods is that the reconstructed function does not satisfy the original line integral data. Kuhl *et al.* (1973) solved this problem, in an approach called orthogonal tangent correction, by back projecting only two orthogonal scans at a time and correcting the solution at each step so it satisfies all the scans used in previous steps. A second disadvantage of the back projection method is that it produces a star-shaped pattern, shown in Fig. 3, which arises when the back projections of two or more scans overlap in regions outside the domain D. This overlapping outside D makes it difficult to ascertain the actual extent of the reconstructed solution. An even more important limitation of this method is that the function it reconstructs is actually a blurred version of the true solution. Remedies for the last two problems will be discussed below.

Fourier transform method

Due to the nature of the operator L, the one-dimensional Fourier transform of a line integral scan $p(t,\hat{\xi})$ of f is equivalent to the corre-

sponding central section of the two-dimensional Fourier transform of f. The central section referred to is the line passing through the origin and making an angle θ with the horizontal axis in the transform plane. Let the one-dimensional transform of $p(t,\hat{\xi})$ be defined as

$$P(T,\hat{\xi}) \equiv \int_{-\infty}^{\infty} p(t,\hat{\xi}) \exp(-2\pi itT)\, dt \tag{15}$$

$$= \int_{-\infty}^{\infty} \exp(-2\pi itT)\, dt \int_{D} \delta(\underline{r}\cdot\hat{\xi} - t)\, f(\underline{r})\, d^2r \tag{16}$$

Interchanging the order of integration and performing the integration over t results in

$$P(T,\hat{\xi}) = \int_{D} f(\underline{r}) \exp(2\pi i \underline{r}\cdot\hat{\xi}T) d^2r\ . \tag{17}$$

If the two-dimensional transform of f is defined as

$$F(\underline{R}) \equiv \int_{D} f(\underline{r}) \exp(2\pi i \underline{r}\cdot\underline{R}) d^2r \tag{18}$$

Eq. (17) may be restated as

$$P(T,\hat{\xi}) = F(T\hat{\xi}). \tag{19}$$

In the Fourier transform method $F(T\hat{\xi})$ is constructed by transforming the line integral scans and the solution is obtained by inverting $F(T\hat{\xi})$ (Crowther *et al.* 1970). If scans at all angles $\theta\epsilon[0,2\pi)$ are available the Fourier transform of f is completely specified and the solution f is unique. Even when the set of scans is finite this method can provide

a good approximation to f. As discussed in the last section the principal solution of Bracewell and Roberts (1954) may be determined uniquely by a discrete sampling of its Fourier transform. If the number of scans is large enough to permit an adequate sampling of the transforms $P(T,\hat{\xi})$ this principal solution, generally a good approximation to the true solution, is recovered uniquely. Bracewell and Riddle (1967) showed that in order to reconstruct an object of diameter D_o there should be $N = \pi D_o R_o$ equally spaced scans. Other authors (Gilbert 1972b, Logan 1975) have also discussed this trade-off between the resolution $d = R_o^{-1}$ and the number of scans. DeRosier and Klug (1968) and Vainshtein (1970) noted that symmetry considerations reduce the number of scans needed to obtain a certain resolution. Bracewell and Roberts (1954) suggested possible modifications to be used when the amplitude is not small near the cut-off frequency although such corrections reduce the resolution of the solution. Although the Fourier transform method produces good reconstructions, two-dimensional inversions are computationally very time consuming; this is the primary drawback of this method. If $F(\underset{\sim}{R})$ is expressed in polar coordinates a one-dimensional Fourier-Bessel inversion may be used (Crowther *et al.* 1970); this inversion requires less computing time but reduces the resolution.

Gilbert (1972b) developed a reconstruction method based on the equivalence (19) which does not require a two-dimensional inversion. This method makes use of the convolution theorem for Fourier transforms which states that the Fourier transform of the convolution of two functions is the product of their transforms (Bracewell 1965, Ch. 6). If the transform $\hat{F}(T\hat{\xi})$ used in a numerical inversion is regarded as the product of the continuous transform and a sampling function $S(T,\theta)$,

$$\hat{F}(T\hat{\xi}) = F(T\hat{\xi})\ S(T,\theta), \tag{20}$$

the solution $\hat{f}(\underset{\sim}{r})$ obtained by inverting $\hat{F}(T\hat{\xi})$ is

$$\hat{f}(\underset{\sim}{r}) = f(\underset{\sim}{r}) * s(t,\theta) = p(t,\hat{\xi}) * s(t,\theta) = \int_{-1}^{1} p(t',\hat{\xi})\ s(t-t',\theta)\ dt' \tag{21}$$

where $s(t,\theta)$ is the inverse transform of $S(T,\theta)$. The goal is to choose the sampling function so its inverse transform is near unity in the domain D. One possible sampling function is the spoke-like function

$$S(T,\theta) \equiv \begin{cases} \delta(X\cos\theta + Y\sin\theta - T) & T \leq 1 \\ 0 & T > 1 \end{cases} \qquad \begin{matrix} \theta = 2\pi\ j/N \\ j=1,\ \ldots,\ N \end{matrix} \tag{22}$$

which samples the transform with unit weight at the origin and along equally spaced radial lines corresponding to the angles of the scans. Using this sampling function Gilbert derived the following relation between the function reconstructed using the back projection method (Vainshtein 1970) and the true solution:

$$f_{bp}(\underset{\sim}{r}) = f(\underset{\sim}{r}) * s(t,\theta). \tag{23}$$

In the limit as $N\to\infty$

$$f_{bp}(\underset{\sim}{r}) = f(\underset{\sim}{r}) * \frac{1}{2\pi t}, \tag{24}$$

a result also obtained by Vainshtein (1971). This formulation illustrates the primary limitation of the back projection method, that the reconstructed function is in fact the actual solution convolved with t^{-1}. In terms of Gilbert's method this is a result of sampling the transform plane with equal weight at all radii without compensating

for the overlapping of the scans near t=0. The sampling function $S'(T,\theta) = T\,S(T,\theta)$ produces better results. Since the error introduced by the convolution of the solution with $S'(T,\theta)$ is of the same order as the truncation error in a two-dimensional Fourier inversion the two approaches are equivalent.

Weighting by the transform variable T was also proposed by Bates and Peters (1971) in the development of rho-filtered layergrams. They formed what they called a layergram by averaging all the line integral scans. They then took the two-dimensional Fourier transform of this average, weighted the transform by T (ρ in their notation), and inverted the modified transform to get the solution. They circumvented the slowness of a two-dimensional numerical inversion by using a system of convex lenses to compute the transforms optically.

Convolution method

The convolution method also depends on the convolution theorem for Fourier transforms. In Gilbert's method the sum of the line integral scans is expressed as the convolution of the desired solution f with the inverse transform of a sampling function which is chosen so that its inverse transform is near unity in the domain D. In the convolution or filtered back projection method the scans are first convolved with a filter function and then summed to yield f. The basis for the convolution technique is the transformation of F(X,Y) from Cartesian coordinates X,Y to polar coordinates T,θ which results in

$$f(\underset{\sim}{r}) = \int_0^{\pi} d\theta \int_{-\infty}^{\infty} F(T\hat{\xi})\ |T|\ \exp(2\pi i t T)\ dT \qquad (25)$$

where $t = x\cos\theta + y\sin\theta$ and $|T|$ is the Jacobian of the transformation.

The equivalence $F(T\hat{\xi}) = P(T,\hat{\xi})$ allows this to be rewritten as

$$f(\underset{\sim}{r}) = \int_0^{\pi} d\theta \int_{-\infty}^{\infty} P(T,\hat{\xi})\ |T|\ \exp(2\pi i t T)\ dT \tag{26}$$

which is, by virtue of the convolution theorem,

$$f(\underset{\sim}{r}) = \int_0^{\pi} p * h\ d\theta = \int_0^{\pi} d\theta \int_{-\infty}^{\infty} p(t',\hat{\xi})\ h(t-t')\ dt'. \tag{27}$$

The filter function h(t) is chosen so its Fourier transform is

$$H(T) \simeq |T|. \tag{28}$$

Convolution of a scan with such a filter amounts to weighting the value of the integral along a particular line $L(t,\theta)$ based on the distance of that line from the point $\underset{\sim}{r}$ at which the function is reconstructed. The contributions of integrals along lines near $\underset{\sim}{r}$ are emphasized whereas those at a distance are of less importance. Generally a decrease in the amplitude as t^{-2} is sufficient to provide a good reconstruction. The same filter function is convolved with the scan at each angle.

Explicit inversion of Eq. (28) to obtain h(t) is not possible since the integral

$$h(t) \simeq \int_{-\infty}^{\infty} |T|\ \exp(-2\pi i t T)\ dT \tag{29}$$

diverges; therefore an approximation of the integral is necessary. One possible approximation for the integral comes from limiting the range of integration to a finite interval $[-T_o, T_o]$. The resulting integral will converge numerically if T_o is not too large. Ramachandran and

Lakshminarayanan (1971) evaluated this finite integral for discrete values of t separated by a distance b; for this they chose $T_o = (2b)^{-1}$. This discretization reduces the inner integral in Eq. (27) to a finite sum. To decrease the computation time further they assumed h(t) to be linear between the points t=kb, k=0,±1, ..., $\pm b^{-1}$. This assumption allows a linear interpolation between the values t = kb at which the inner sum in Eq. (27) is evaluated and the values $t = x\cos\theta + y\sin\theta$ needed for the reconstruction of f. A second approximation for the integral (29) is based on the fact that $F(T\hat{\xi})$, and hence $P(T,\hat{\xi})$, is small for $|T| > 1$. As a result the requirement that $H(T) \simeq |T|$ may be weakened for $|T| > 1$ without greatly altering the inner integral in Eq. (26). T_o therefore may be taken to be 1 and the choice

$$H(T) = \begin{cases} |T| & T < 1 \\ 0 & T \geq 1 \end{cases} \tag{30}$$

is acceptable. The reconstruction obtained using a continuous filter defined in this way (Bracewell and Riddle 1967) is poor because h(t) decays only as t^{-1}. The piecewise linear filter of Ramachandran and Lakshminarayanan which decays as t^{-2} produces better reconstructions. Shepp and Logan (1974) showed that the assumption of piecewise linearity is consistent with the observation that $H(T) \simeq |T|$ for $|T| < 1$ and therefore does not imply a loss of accuracy. All these filters have a large positive amplitude over an interval centered at t=0 which is flanked by negative or alternately negative and positive side lobes with progressively smaller amplitudes. The negative side lobes tend to eliminate the star pattern characteristic of the back projection method, as illustrated in Fig. 4. In this diagram the solid lines mark the extent of the positive

central lobes of p_3' and p_4'. The heavy shading shows regions in which this overlapping is offset by negative contributions from the back projections of p_1' and p_2'. This reduction in the intensity of the overlap makes it easier to define the reconstructed solution.

The approximate solution $\hat{f}$ obtained using the convolution method described above can be thought of as a blurred version of the true solution f,

$$\hat{f} = f * \psi. \tag{31}$$

Davison and Grünbaum (1979) chose a set of filter functions $\{\alpha_i\}_{i=1}^{N}$ so that the reconstructed function

$$\hat{f} = f * \frac{1}{N} \sum_{i=1}^{N} B_i \alpha_i , \tag{32}$$

where the back projection operator B_i is defined by $(B_i\alpha_i)(x,y) = \alpha_i(\underset{\sim}{r}\cdot\hat{\xi}_i)$, is close to $f * \phi$ for a known ϕ. The point response function ϕ is chosen to be some sort of identity function which determines the resolution of the reconstruction. This approach does not require equally spaced angles as do the convolution algorithms mentioned previously. Cho *et al.* (1974) attempted to reduce the blurring by using a filter based on a geometrical correction of the projection data in the spatial domain.

Series expansion methods

In the series expansion approach to the tomography problem the unknown function is expressed as a sum of known orthogonal basis functions $\{b_i(\underset{\sim}{r})\}$,

$$f(\underset{\sim}{r}) = \sum_i a_i b_i(\underset{\sim}{r}). \tag{33}$$

Applying the operator L results in

$$p(t,\hat{\xi}) = \sum_i a_i \int_{L(t,\theta)} b_i(\underset{\sim}{r})\, ds. \tag{34}$$

The problem is then reduced to finding the unknown coefficients a_i. If the set of basis functions is complete the reconstructed solution will be unique. Cormack (1963) expanded the solution expressed in polar coordinates, r,β as an infinite Fourier series

$$f(\underset{\sim}{r}) = \sum_{n=-\infty}^{\infty} f_n(r)\, \exp(in\beta) \tag{35}$$

where

$$f_n(r) = \frac{1}{2\pi}\int_0^{2\pi} f(\underset{\sim}{r})\, \exp(-in\beta)\, d\beta\ . \tag{36}$$

Under the assumption that the f_n's are bounded and piecewise continuous, a similar expansion of the line integral scans $p(t,\hat{\xi})$ results in a unique relationship between the Fourier components of each series. Since the set of Fourier components is complete the reconstructed solution is unique. In a later paper, Cormack (1964) extended these results to include the entire plane and considered alternative expansions in other orthogonal functions. For a set of discrete angles θ_j, $j=1, \ldots, N$, Eq. (34) is

$$p_j(t) = \sum_i a_i\, B_{ij}(t) \qquad j = 1, \ldots, N \tag{37}$$

where

$$p_j(t) = p(t,\hat{\xi}_j) \text{ and } B_{ij}(t) = \int_{L(t,\theta_j)} b_i(\underset{\sim}{r})\, ds.$$

Series expansion solutions are especially useful in obtaining consistency conditions which must be satisfied if a set of functions is to represent the line integral scans of some function. Logan and Shepp (1975) showed that the minimum norm solution to the discrete problem when the angles are equally spaced is of the form

$$f(\underset{\sim}{r}) = \sum_{j=1}^{N} \rho_j(r) \tag{38}$$

where $\rho_j(r) = \rho(\underset{\sim}{r}\cdot\hat{\xi}_j)$ is the ridge function in the direction θ_j. They sought a solution by expanding ρ_j as an infinite Fourier sine series. By similarly expanding the line integral scans of ρ_j they obtained certain consistency conditions which must be satisfied if a set of functions $\{g(t,\hat{\xi}_j)\}_{j=1}^{N}$ is to correspond to a unique minimum norm solution. These conditions are that: 1) $g(t,\hat{\xi}_j) = 0$ for $|t| > 1$; 2) the integral $\int_{-1}^{1} [g(t,\hat{\xi}_j)]^2 (1-t^2)^{-1/2}\, dt$ is finite; and 3) the Fourier components of $g(t,\hat{\xi}_j)$ satisfy a certain relationship. They have incorporated the symmetry condition $g(t,\hat{\xi}_j) = g(-t,-\hat{\xi}_j)$ by restricting the range of angles to $[0,\pi)$. Marr (1974) looked for a least squares polynomial solution to the fan beam tomography problem. The consistency conditions imposed by his finite Fourier series expansion are that the data be symmetric and orthogonal to $\exp(in\theta)$. Ludwig (1966) showed that the consistency conditions implied by an expansion in spherical harmonics are equivalent to those specified above for the general solution to the tomography problem. In a preliminary report, Miller (1978) proposed a partial eigenfunction expansion.

Iterative methods

The series expansion formulation (37) is a starting point for the

iterative and matrix inversion reconstruction techniques. The domain D is enclosed by a qXq grid as shown in Fig. 5 and the solution $f(\underset{\sim}{r})$ is assumed to have a constant mean value f_i in the element R_i, i=1, ..., q^2. This amounts to choosing basis functions

$$b_i(\underset{\sim}{r}) = \begin{cases} 1 & r \epsilon R_i \\ 0 & r \notin R_i \end{cases} \qquad i = 1, \ldots, q^2. \tag{39}$$

The unknown coefficients a_i in Eq. (37) are then the f_i's. The projection data are regarded as sets of K integrals over strips S_{jk}, k=1, ..., K, where the first subscript denotes the projection angle θ_j, j=1, ..., N. The total number of data is then KN and the projections may be renumbered to form the set $\{p_j\}_{j=1}^{KN}$. The projection operation is then

$$p_j = \sum_{i=1}^{q^2} w_{ij} f_i \qquad j = 1, \ldots, KN \tag{40}$$

where the weight function w_{ij} represents the fraction of the i^{th} element traversed by the strip S_{jk}. Since each strip intersects only a few of the elements R_i the matrix W composed of the elements w_{ij} is quite sparse.

There are three main iterative methods which have been applied to the tomography problem. They differ in the choice of a correction Δf_i^{ℓ} applied to f_i^{ℓ}, the ℓ^{th} iterate of the solution, to obtain

$$f_i^{\ell+1} = f_i^{\ell} + \Delta f_i^{\ell}$$

or $\qquad i = 1, \ldots, q^2. \tag{41}$

$$f_i^{\ell+1} = \Delta f_i^{\ell} \cdot f_i^{\ell}$$

In the Iterative Least Squares Technique (ILST) (Budinger and Gullberg 1974b) the additive correction Δf_i^{ℓ} is chosen so as to minimize the

squared error between the set of projections of the solution f^{ℓ} composed of all elements f_i^{ℓ}, i=1, ..., q^2, and the observed projections. The Algebraic Reconstruction Technique (ART) (Gordon *et al.* 1970) uses a projection p_j to correct the estimates f_i for elements R_i which intersect the strip $S_{j'k}$ corresponding to p_j. These corrected estimates are then used in the calculation of corrections using subsequent projections. Hamaker and Solomon (1978) gave a rate of convergence for this method which depends on the projection angles. The Simultaneous Iterative Reconstruction Technique (SIRT) (Gilbert 1972a) uses all the projections to correct the solution in each element R_i. ART and SIRT have both additive and multiplicative formulations. In both methods the additive formulation includes a positivity constraint on the f_i's. The multiplicative methods will produce positive f_i's if the initial guess is positive.

Singular Value Decomposition

Singular value decomposition is a technique for finding a solution $f(\underset{\sim}{r})$ which is related to a set of data $\{\gamma(\hat{\xi}_j)\}_{j=1}^{N}$ by

$$\gamma(\hat{\xi}_j) = \int_V G(\underset{\sim}{r},\hat{\xi}_j)\, f(\underset{\sim}{r})\, \underline{dr} \qquad j = 1, \ldots, N \tag{42}$$

where the kernel $G(r,\hat{\xi}_j)$ describing the relevant physical process in the domain V is non-symmetric. In the notation of the tomography problem the discrete form of this relation is

$$\underset{\sim}{W}\,\underset{\sim}{f} = \underset{\sim}{p} \tag{43}$$

where $\underset{\sim}{f}$ is a vector composed of the $q^2 = m$ model parameters f_i, $\underset{\sim}{p}$ is the data vector composed of the KN = n projections p_j, and $\underset{\sim}{W}$ is the nXm matrix whose elements are the weighting factors w_{ij}. If $\underset{\sim}{W}$ were non-singular,

symmetric and square its inverse $\underset{\sim}{W}^{-1}$ would exist and the solution to Eq. (43) would be

$$\underset{\sim}{f} = \underset{\sim}{W}^{-1} \underset{\sim}{p}. \tag{44}$$

In general the system is highly underdetermined, n << m, and $\underset{\sim}{W}^{-1}$ does not exist. The object of the singular value decomposition method is the construction of a "pseudo-inverse" for a rectangular or singular square matrix. The following discussion is based on that of Lanczos (1961).

The eigenvalue problem for the symmetric matrix $S = \begin{pmatrix} \underset{\sim}{O} & \underset{\sim}{W} \\ \underset{\sim}{W}^T & \underset{\sim}{O} \end{pmatrix}$ reduces to

$$\begin{aligned} \underset{\sim}{W}\, \underset{\sim}{v}_j &= \lambda_j\, \underset{\sim}{u}_j \qquad j = 1, \ldots, m \\ \underset{\sim}{W}^T \underset{\sim}{u}_i &= \lambda_i\, \underset{\sim}{v}_i \qquad i = 1, \ldots, n \end{aligned} \tag{45}$$

or equivalently

$$\begin{aligned} \underset{\sim}{W}\, \underset{\sim}{W}^T\, \underset{\sim}{u}_i &= \lambda_i^2\, \underset{\sim}{u}_i \qquad i = 1, \ldots, n \\ \underset{\sim}{W}^T\, \underset{\sim}{W}\, \underset{\sim}{v}_j &= \lambda_j^2\, \underset{\sim}{v}_j \qquad j = 1, \ldots, m \end{aligned} \tag{46}$$

where $\lambda_i = \lambda_j$ for i=j. The values λ_j^2 are eigenvalues of the symmetric matrices $\underset{\sim}{W}\, \underset{\sim}{W}^T$ and $\underset{\sim}{W}^T\, \underset{\sim}{W}$; the λ_j's are known as singular values of $\underset{\sim}{W}$. Let ℓ be the number of non-zero eigenvalues. Then $\lambda_j = 0$ for $j > \ell$ if the singular values are numbered in order of decreasing magnitude. Let $\underset{\sim}{U}$ and $\underset{\sim}{V}$ be unitary matrices whose columns are, respectively, the eigenvectors $\underset{\sim}{u}_i$ and $\underset{\sim}{v}_j$ associated with the ℓ non-zero eigenvalues. $\underset{\sim}{U}$ is n$X\ell$ and $\underset{\sim}{V}$ is m$X\ell$. Let $\underset{\sim}{U}_o$ and $\underset{\sim}{V}_o$ denote, respectively, the nX(n-ℓ) and mX(m-ℓ) unitary matrices composed of the remaining eigenvectors. The matrix $\underset{\sim}{W}$ may be decomposed into

$$\underset{\sim}{W} = \underset{\sim}{U}\, \underset{\sim}{\Omega}\, \underset{\sim}{V}^T \tag{47}$$

where $\underset{\sim}{\Omega}$ is a diagonal matrix whose elements are the singular values λ_j, $j = 1, \ldots, \ell$. Although the operator $\underset{\sim}{W}$ has no effect on elements of $\underset{\sim}{U}_o$ and $\underset{\sim}{V}_o$ these subspaces are important in establishing the existence and uniqueness of solutions to $\underset{\sim}{W}\underset{\sim}{f} = \underset{\sim}{p}$. If $\ell < n$, a solution exists if and only if the data vector $\underset{\sim}{p}$ is orthogonal to all solutions of the homogeneous equation $\underset{\sim}{W}\,\underset{\sim}{u}_i = 0$, that is, if $\underset{\sim}{U}_0^T\,\underset{\sim}{b} = 0$. If $\ell = n$, then $\underset{\sim}{U}_o \equiv 0$ and a solution will always exist. If $n < \ell < m$, there will be $m - \ell$ solutions; a unique solution is possible only if $\ell = m$, i.e., $\underset{\sim}{V}_o \equiv 0$.

The natural inverse or the generalized inverse (Penrose 1955) of $\underset{\sim}{W}$ is

$$\underset{\sim}{W}^+ = \underset{\sim}{V}\,\underset{\sim}{\Omega}^{-1}\,\underset{\sim}{U}^T \tag{48}$$

where $\underset{\sim}{\Omega}^{-1}$ is the diagonal matrix composed of the elements λ_j^{-1}, $j = 1, \ldots, \ell$. The product $\underset{\sim}{W}\,\underset{\sim}{W}^+ = \underset{\sim}{U}\,\underset{\sim}{U}^T$ may be thought of as an identity for the subspace $\underset{\sim}{U}$ in the sense that multiplication of any $\underset{\sim}{u}\varepsilon\underset{\sim}{U}$ by $\underset{\sim}{W}\,\underset{\sim}{W}^+$ reproduces $\underset{\sim}{u}$. If $\underset{\sim}{W}\,\underset{\sim}{W}^+ = \underset{\sim}{I}_n$, then $\underset{\sim}{U}_o \equiv 0$ and the consistency condition $\underset{\sim}{U}_o^T\,\underset{\sim}{b} = 0$ is identically satisfied. In this case there is a solution corresponding to any data vector $\underset{\sim}{b}$. The product $\underset{\sim}{W}^+\,\underset{\sim}{W}$ may be considered an identity for $\underset{\sim}{V}$ in a similar manner. If $\underset{\sim}{W}^+\,\underset{\sim}{W} = \underset{\sim}{I}_m$, then $\underset{\sim}{V}_o \equiv 0$ and if a solution exists it is unique. The solution

$$\hat{\underset{\sim}{f}} = \underset{\sim}{W}^+\,\underset{\sim}{p} \tag{49}$$

is seen by substitution to satisfy Eq. (43) for $\underset{\sim}{p}\varepsilon\underset{\sim}{U}$. Thus $\underset{\sim}{p}$ is orthogonal to all $\underset{\sim}{u}\varepsilon\underset{\sim}{U}_o$ and satisfies the compatability condition for the existence of a solution. Penrose (1956) shows that the solution (49) is the best least squares solution to Eq. (43) in the sense that

$$||\underset{\sim}{W}\,\underset{\sim}{f} - \underset{\sim}{p}|| > ||\underset{\sim}{W}\,\hat{\underset{\sim}{f}} - \underset{\sim}{p}|| \tag{50}$$

or

$$||\underset{\sim}{W}\,\underset{\sim}{f} - \underset{\sim}{p}|| = ||\underset{\sim}{W}\,\hat{\underset{\sim}{f}} - \underset{\sim}{p}|| \text{ and } ||f|| \geq ||\hat{f}||, \tag{51}$$

where $||x|| = \underset{\sim}{x}^T\underset{\sim}{x}$, for any other solution f.

Budinger and Gullberg (1974a) used the generalized inverse to find a least squares solution to the tomography problem. They incorporated the presence of errors in the data by minimizing instead of $\underset{\sim}{\varepsilon}^T\underset{\sim}{\varepsilon}$, where $\underset{\sim}{\varepsilon} = \underset{\sim}{W}\underset{\sim}{f} - \underset{\sim}{p}$, the quantity $\underset{\sim}{\varepsilon}^T\underset{\sim}{P}\underset{\sim}{\varepsilon}$ where $\underset{\sim}{P}$ is the inverse covariance matrix for the data. They gave the solution

$$\underset{\sim}{f} = (\underset{\sim}{W}^T\,\underset{\sim}{P}\,\underset{\sim}{W})^{-1}\,\underset{\sim}{W}^T\,\underset{\sim}{P}\,\underset{\sim}{p} \tag{52}$$

or if $\underset{\sim}{W}^T\,\underset{\sim}{P}\,\underset{\sim}{W}$ is singular, as is usually the case,

$$\underset{\sim}{f} = (\underset{\sim}{P}^{\frac{1}{2}}\,\underset{\sim}{W})^{+}\,\underset{\sim}{P}^{\frac{1}{2}}\,\underset{\sim}{p}. \tag{53}$$

Tewarson (1972) minimized the quadratic form

$$\phi(\underset{\sim}{f}) = \underset{\sim}{\varepsilon}^T\underset{\sim}{P}\underset{\sim}{\varepsilon} + \underset{\sim}{f}^T\underset{\sim}{F}\underset{\sim}{f} \tag{54}$$

with the solution

$$\underset{\sim}{f} = (\underset{\sim}{W}^T\underset{\sim}{P}\underset{\sim}{W} + \underset{\sim}{F})^{+}\,\underset{\sim}{W}^T\underset{\sim}{P}\underset{\sim}{p}. \tag{55}$$

The matrices $\underset{\sim}{P}$ and $\underset{\sim}{F}$ may be chosen to be the inverses of the covariance matrices of the data and the model parameters, respectively, or they may be determined using the regularization method of Tikhonov (Tikhonov and Arsenin 1977).

Evaluation of reconstruction methods

Reconstructions based on projections at a finite number of angles

cannot in general reproduce exactly the source function and approximate methods for reconstruction are generally judged by the closeness of the reconstructed function to the original function. One technique for evaluating this accuracy is to measure a set of projections of a physical object with known properties, called a phantom, and compare the object reconstructed using these projections with the original object. This evaluation approach is deficient in that the inaccuracies which result may be due not to the reconstruction method but to experimental and computational limitations. Experimental errors are eliminated by using instead a mathematical phantom. A set of projections of a known function is calculated and the resulting function compared with the source function. Since reconstructive tomography is by its very nature a computational phenomenon, inaccuracies due to round-off and other purely numerical errors cannot be separated from inaccuracies due solely to the reconstruction method itself. The accuracy of some methods may be characterized in terms of an error bound on the reconstructed function or, in the iterative methods, a rate of convergence of the algorithm. Gordon (1974) suggested an additional assessment of the accuracy of reconstructions using iterative techniques. The presence of some feature in a reconstructed solution may be tested by removing the feature and using the altered solution as an initial guess for a second application of the reconstruction procedure. If the feature appears again in the second or subsequent reconstruction it is not likely to be an artifact of the reconstruction method.

CHAPTER III

STABILITY OF SOLUTIONS

As described in the introduction an inverse problem is well posed if, given any data set, a solution exists, is unique, and depends continuously on the data. For the tomography problem Radon (1917) proved that the first two requirements are met by a solution f subject to the restriction that the integral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f(\underset{\sim}{r})}{r} \, d^2 r$$

be finite, but he did not examine the third criterion. Ludwig (1966) showed that the inversion is indeed stable within the space of rapidly decreasing functions, that is, those functions $f(\underset{\sim}{r})$ for which derivatives of all orders exist, are continuous, and vanish more rapidly than r^k, $k > 0$, as $r \to \infty$. In general, however, the problem of inverting the line integral data to obtain f is not well-posed.

The stability of the discrete problem

$$\underset{\sim}{W} \, \underset{\sim}{f} = \underset{\sim}{p} \tag{56}$$

with a finite data set may be expressed in terms of the singular values of the matrix W. Since the problem is linear, perturbations $\delta\underset{\sim}{f}$ and $\delta\underset{\sim}{p}$, not necessarily small, in, respectively, the solution and the data are related by

$$\underset{\sim}{W} \, \delta\underset{\sim}{f} = \delta\underset{\sim}{p} \tag{57}$$

(Lanczos 1961). The decomposition $\underset{\sim}{W} = \underset{\sim}{U}\underset{\sim}{\Omega}\underset{\sim}{V}^T$ suggests the transformation $\delta\underset{\sim}{p} = \underset{\sim}{U}\delta\underset{\sim}{p}'$, $\delta\underset{\sim}{f} = \underset{\sim}{V}\delta\underset{\sim}{f}'$ which results in

$$\underset{\sim}{\Omega}\ \delta\underset{\sim}{f}' = \delta\underset{\sim}{p}'. \tag{58}$$

Since $\underset{\sim}{U}$ and $\underset{\sim}{V}$ are unitary matrices $|\delta\underset{\sim}{f}'| = |\delta\underset{\sim}{f}|$ and $|\delta\underset{\sim}{p}'| = |\delta\underset{\sim}{p}|$ and the variation in the ith component of the solution is

$$\delta f_i = \lambda_i^{-1}\ \delta p_i\ . \tag{59}$$

Thus a small error in a data value associated with a small singular value can cause an unacceptably large variation in the solution.

A more quantitative measure of the stability of a discrete inverse problem is the variance of the solution. For an inverse problem to be stable the variance should be small. The variance of the kth component of the solution to $\underset{\sim}{W}\ \underset{\sim}{f} = \underset{\sim}{p}$ which minimizes $\underset{\sim}{f}^T\underset{\sim}{f}$ is

$$\text{var}\ (f_k) = \sigma^2 \sum_{i=1}^{\ell} \frac{v_{ki}^2}{\lambda_i^2} \tag{60}$$

where ℓ is the number of non-zero singular values and σ^2 is the uncertainty in the data due to measurement errors (Munk and Wunsch 1979). In this formulation the errors in the data are assumed to be uncorrelated. If the errors in the data are indeed correlated this formulation is still valid provided certain transformations involving the covariance matrices for the data and the model parameters are made (Jackson 1972, Munk and Wunsch 1979). If one or more of the singular values are small the variance will be very large. One way to reduce the variance is to exclude very small singular values by establishing a threshold below which the

singular values are considered zero. The matrix $\underline{V}$ would then be composed of the eigenvectors corresponding to singular values greater than the threshold value. The threshold may be chosen so the uncertainty in the solution is sufficiently small. Unfortunately, decreasing the number of eigenvectors in $\underline{V}$ also decreases the resolution in the solution. A unique, perfectly resolved solution is possible only when none of the singular values vanish. The most appropriate compromise in this trade-off between minimizing the variance and maximizing the resolution must be decided for each problem. Davison and Grünbaum (1979) noted a similar trade-off between the spatial resolution achieved using a particular convolution algorithm and the closeness of the point response function for that algorithm to a specified point response function.

Smith *et al.* (1977) approached the problem of errors in the data by examining whether a set of noisy line integrals $p(t,\hat{\xi}_j) = p_j(t)$, $j=1, \ldots, N$, can actually represent the line integrals of some function f. They proved that if the numbers

$$\alpha_{jk} = \int_{-1}^{1} p_j(t)\ t^k\ dt \tag{61}$$

are the values at the points $\hat{\xi}_j$ of a homogeneous polynomial of degree $k \leq N-2$, then there does exist a function f whose line integral scans in the directions θ_j are the given data $p_j(t)$. This conclusion is a result of the form of the operators L_j, as expressed by the equivalence (11).

CHAPTER IV

PROSPECTS FOR FUTURE RESEARCH

Much progress has been made on the tomography problem, but some work remains to be done. One major possibility for future research is the incorporation of *a priori* information into the problem. There are several ways in which this may be done. One way which was discussed previously is to use knowledge of the symmetry properties of the function to reduce the number of scans needed for reconstruction. DeRosier and Klug (1968) discussed the number needed in electron microscopy to reconstruct molecules with different types of symmetry. A second type of *a priori* information which may be useful is a bound on the unknown function. Miller (1978) proposed using the boundedness of the norm of f as a stabilizing constraint. Davison and Grünbaum (1979) used this bound to establish an error bound for their convolution technique. A third possibility is the incorporation of *a priori* knowledge into the construction of the covariance matrix for the solution (Munk and Wunsch 1979).

Constraints such as the positivity of the unknown function or the total mass of the object may be used in evaluating the validity of a reconstructed solution. A solution which does not satisfy these constraints cannot be a viable solution. So far the positivity constraint has been included only in the iterative methods (Gilbert 1972a, Gordon *et al.* 1970); any component of the solution which turns out to be

negative is set to zero. The problem of solving

$$\underset{\sim}{W}\,\underset{\sim}{f} = \underset{\sim}{p} \tag{62}$$

subject to the constraint

$$f_i > 0 \quad . \tag{63}$$

might also be approached using techniques in linear programming (Sabatier 1977a, 1977b) in which the constraint is incorporated into the formulation of the problem.

REFERENCES

Aki, K., A. Christoffersen and E.S. Husebye (1977) Determination of the three-dimensional seismic structure of the lithosphere. *J. Geophys. Res.* 82:277-296.

Backus, G.E. (1970) Non-uniqueness of the external geomagnetic field determined by surface intensity measurements. *J. Geophys. Res.* 75:6339-6341.

Bates, R.H.T. and T.M. Peters (1971) Towards improvements in tomography. *N.Z.J. Sci.* 14:883-896.

Bracewell, R.N. (1956) Strip integration in radio astronomy. *Australian J. Phys.* 9:198-217.

_____ (1965) *The Fourier Transform and Its Applications.* New York: McGraw-Hill, 381 pp.

_____ and A.C. Riddle (1967) Inversion of fan-beam scans in radio astronomy. *Astrophys. J.* 150:427-434.

_____ and J.A. Roberts (1954) Aerial smoothing in radio astronomy. *Australian J. Phys.* 7:615-640.

Brooks, R.A. and G. DiChiro (1976) Principles of Computer Assisted Tomography (CAT) in radiographic and radioisotope imaging. *Phys. Med. Biol.* 21:689-732.

Budinger, T.F. and G.T. Gullberg (1974a) Three-dimensional reconstruction in nuclear medicine by iterative least squares and Fourier transform techniques. Lawrence Berkeley Laboratory Report LBL-2146.

_____ and G.T. Gullberg (1974b) Three-dimensional reconstruction in nuclear medicine emission imaging. *IEEE Trans. Nucl. Sci.* NS-21:2-20.

Cho, Z.H., I. Ahm, C. Bohm and G. Huth (1974) Computerized image reconstruction methods with multiple photon/x-ray transmission scanning. *Phys. Med. Biol.* 19:511-522.

Cormack, A.M. (1963) Representation of a function by its line integrals, with some radiological applications. *J. Appl. Phys.* 34:2722-2727.

_____ (1964) Representation of a function by its line integrals, with some radiological applications. II. *J. Appl. Phys.* 35:2908-2913.

Crowther, R.A., D.J. DeRosier and A. Klug (1970) The reconstruction of a three-dimensional structure from projections and its applications to electron microscopy. *Proc. Roy. Soc. Lond. A* 317:319-340.

Davison, M.E. and F.A. Grünbaum (1979) Convolution algorithms for arbitrary projection angles. Preprint.

DeRosier, D.J. and A. Klug (1968) Reconstruction of three-dimensional structures from electron micrographs. *Nature* 217:130-134.

Frieder, G. and G.T. Herman (1971) Resolution in reconstructing objects from electron micrographs. *J. Theor. Biol.* 33:189-211.

Gel'fand, I.M., M.I. Graev and N. Ya. Vilenkin (1966) *Generalized Functions, Vol. 5, Integral Geometry and Representation Theory.* Translated by Eugene Saletan. New York: Academic Press, 449 pp.

Gilbert, P.F.C. (1972a) Iterative methods for the three-dimensional reconstruction of an object from projections. *J. Theor. Biol.* 36:105-117.

_____ (1972b) The reconstruction of a three-dimensional structure from projections and its application to electron microscopy-II. Direct methods. *Proc. Roy. Soc. Lond. B* 182:89-102.

Gordon, R. (1974) A tutorial on ART (Algebraic Reconstruction Techniques). *IEEE Trans. Nucl. Sci.* NS-21:78-93.

_____, R. Bender and G.T. Herman (1970) Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and x-ray photography. *J. Theor. Biol.* 29:471-481.

_____ and G.T. Herman (1974) Three-dimensional reconstruction from projections: a review of algorithms. *Int. Rev. Cytol.* 38:111-151.

Hamaker, C. and D.C. Solomon (1978) The angles between the null spaces of x-rays. *J. Math. Anal. Appl.* 62:1-23.

Hart, R.G. (1968) Electron microscopy of unstained biological material: the polytropic montage. *Science* 159:1464-1467.

Herman, G.T. and A. Naparstek (1977) Fast image reconstruction based on a Radon inversion formula appropriate for rapidly collected data. *SIAM J. Appl. Math.* 33:511-533.

Jackson, D.D. (1972) Interpretation of inaccurate, insufficient and inconsistent data. *Geophys. J. R. astr. Soc.* 28:97-109.

Kuhl, D.E. and R.Q. Edwards (1968) Reorganizing data from transverse section scans of the brain using digital processing. *Radiology* 91:975-983.

Kuhl, D.E., R.Q. Edwards, A.R. Ricci and M. Reivich (1973) Quantitative section scanning using orthogonal tangent correction. *J. Nucl. Med.* 14:196-200.

Landzos, C. (1961) *Linear Differential Operators.* London: Van Nostrand, 564 pp.

Logan, B.F. (1975) The uncertainty principle in reconstructing functions from projections. *Duke Math. J.* 42:661-706.

_____ and L.A. Shepp (1975) Optimal reconstruction of a function from its projections. *Duke Math. J.* 42:645-659.

Ludwig, D. (1966) The Radon transform on euclidean space. *Comm. Pure Appl. Math.* 19:49-81.

Marr, R.B. (1974) On the reconstruction of a function on a circular domain from a sampling of its line integrals. *J. Math. Anal. Appl.* 45:357-374.

Miller, K. (1978) An optimal method for the x-ray reconstruction problem. *Amer. Math. Soc. Not.* 25:A-161-A-162.

Munk, W. and C. Wunsch (1979) Ocean acoustic tomography: a scheme for large scale monitoring. *Deep-Sea Res.* 26A:123-161.

Parker, R.L. (1972) Inverse theory with grossly inadequate data. *Geophys. J. R. astr. Soc.* 29:123-138.

_____ (1975) The theory of ideal bodies for gravity interpretation. *Geophys. J. R. astr. Soc.* 42:315-334.

_____ (1977) Understanding inverse theory. *Ann. Rev. Earth Planet. Sci.* 5:35-64.

Penrose, R. (1955) A generalized inverse for matrices. *Proc. Camb. Phil. Soc.* 51:406-413.

_____ (1956) On best approximate solutions of linear matrix equations. *Proc. Camb. Phil. Soc.* 52:17-19.

Radon, J. (1917) Über die Bestimmung von Funktionen durch ihre Integralwerte langs gewisser Mannigfaltizkeiten. *Ber. Verh. Sachs. Akad. Wiss. Leipzig, Math.-Nat. Kl.* 69:262-277.

Ramachandran, G.N. and A.V. Lakshminarayanan (1971) Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms. *Proc. Nat. Acad. Sci. USA* 68:2236-2240.

Sabatier, P.C. (1977a) Positivity constraints in linear inverse problems - I. General theory. *Geophys. J. R. astr. Soc.* 48:415-442.

Sabatier, P.C. (1977b) Positivity constraints in linear inverse problems - II. Applications. *Geophys. J. R. astr. Soc.* 48:443-469.

Shepp, L.A. and J.B. Kruskal (1978) Computerized tomography: the new medical x-ray technology. *Amer. Math. Monthly* 85:420-439.

_____ and B.F. Logan (1974) The Fourier reconstruction of a head section. *IEEE Trans. Nucl. Sci.* NS-21:21-43.

Smith, K.T., D.C. Solomon and S.L. Wagner (1977) Practical and mathematical aspects of the problem of reconstructing objects from radiographs. *Bull. Amer. Math. Soc.* 83:1227-1270.

Solomon, D.C. (1976) The x-ray transform. *J. Math. Anal. Appl.* 56:61-83.

Tewarson, R.P. (1972) Solution of linear equations in remote sensing and picture reconstruction. *Computing* 10:221-230.

Tikhonov, A.N. and V.Y. Arsenin (1977) *Solutions of Ill-Posed Problems.* Washington, D.C.: V.H. Winston and Sons, 258 pp.

Vainshtein, B.K. (1970) Finding the structure of objects from projections. *Kristallografiya* 15:894-902. Translated in *Sov. Phys.-Crystallogr.* 15:781-787 (1971).

_____ (1971) Synthesis of projecting functions. *Dokl. Akad. Nauk SSSR* 196:1072-1075. Translated in *Sov. Phys.-Dokl.* 16:66-69 (1971).

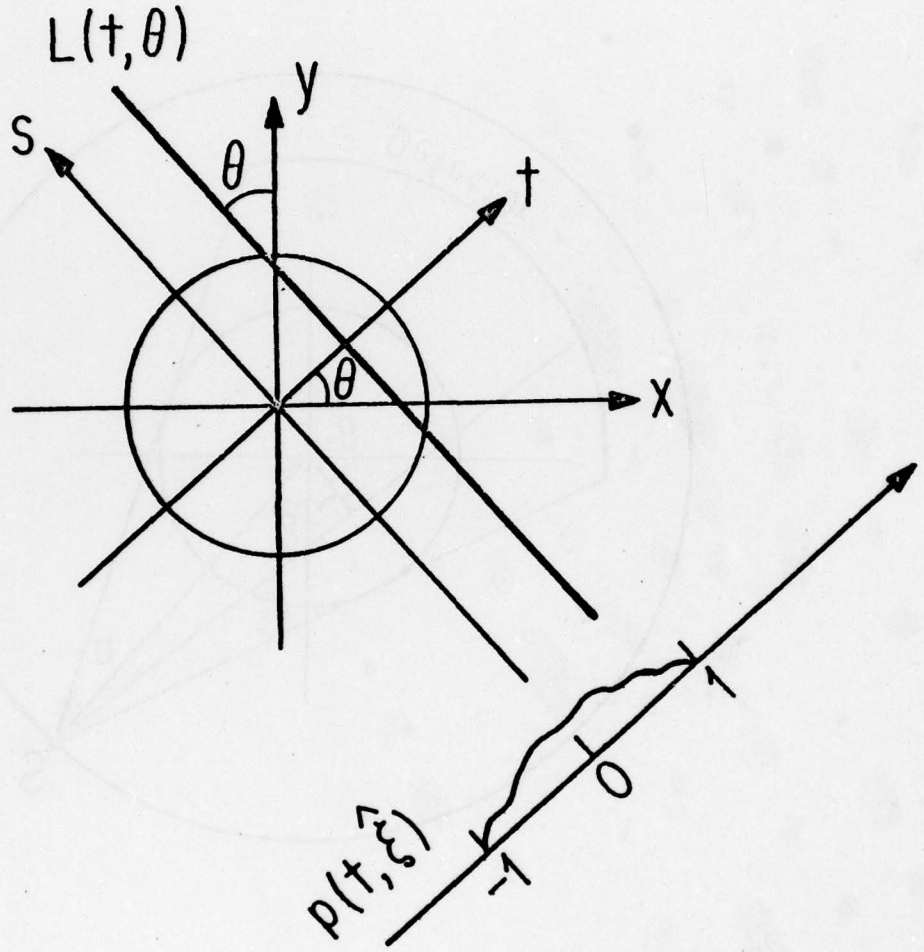


Fig. 1--Geometry of line integral data for parallel mode tomography (after Vainshtein 1970).

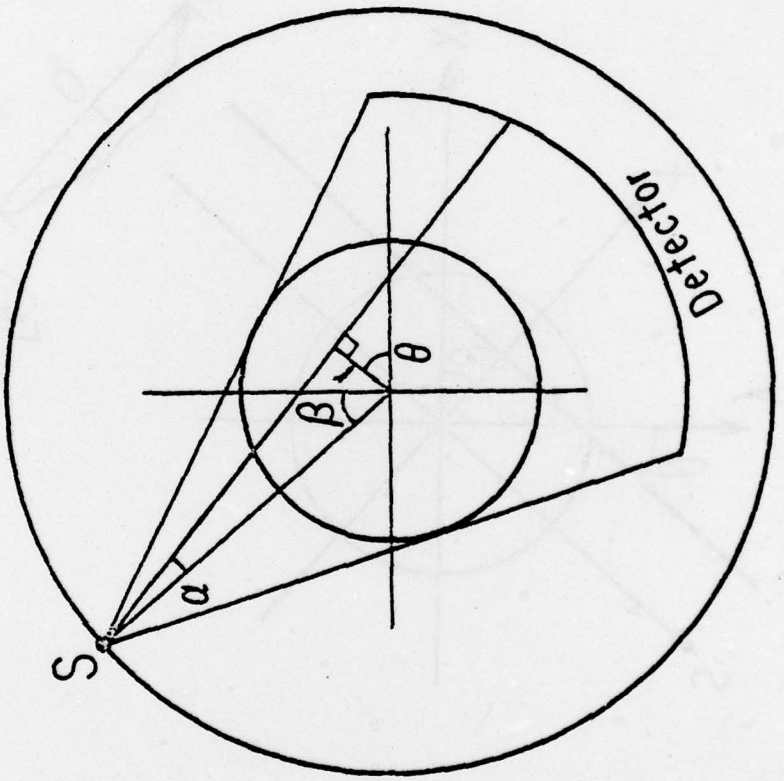


Fig. 2--Geometry of fan beam scanning (after Herman and Naparstek 1977).

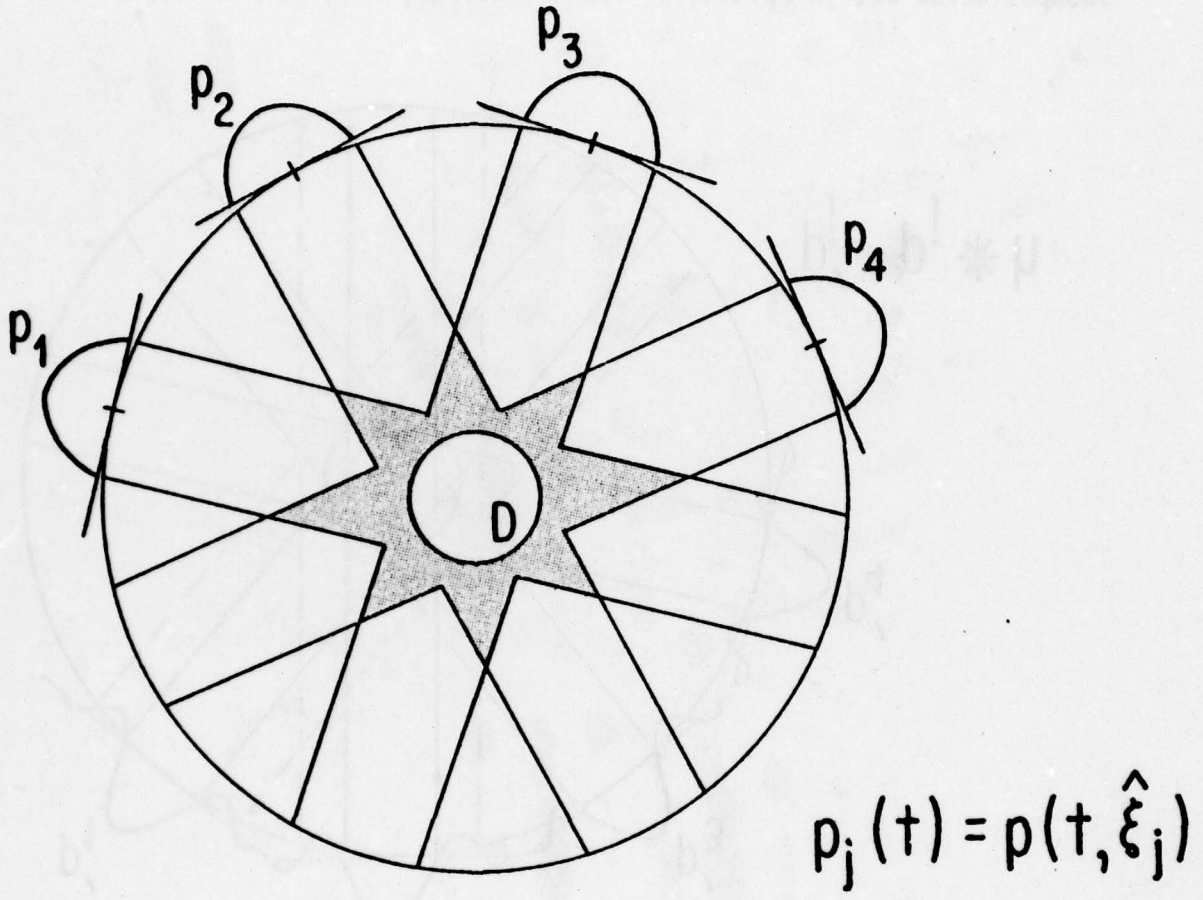


Fig. 3--Star-shaped artifact in back projection method. See text for explanation.

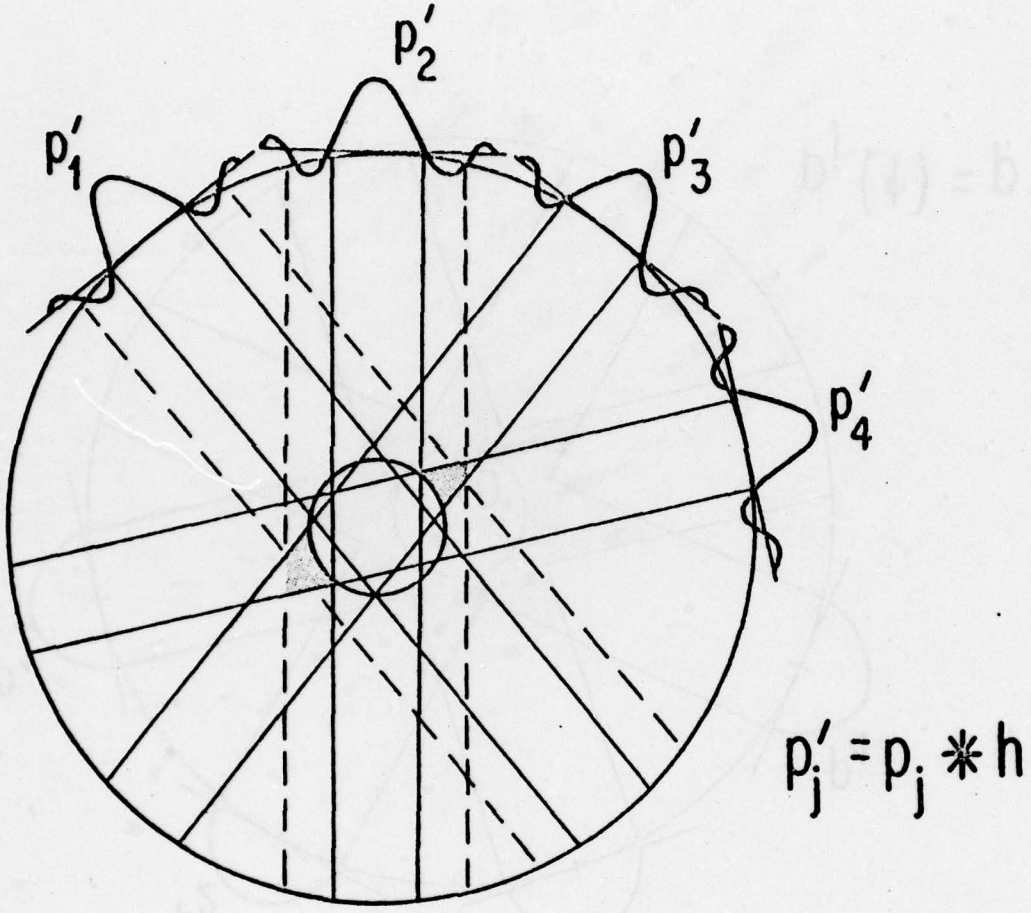


Fig. 4--Convolution or filtered back projection method. Filtering of the scans reduces the star artifact produced by the back projection method. See text for explanation.

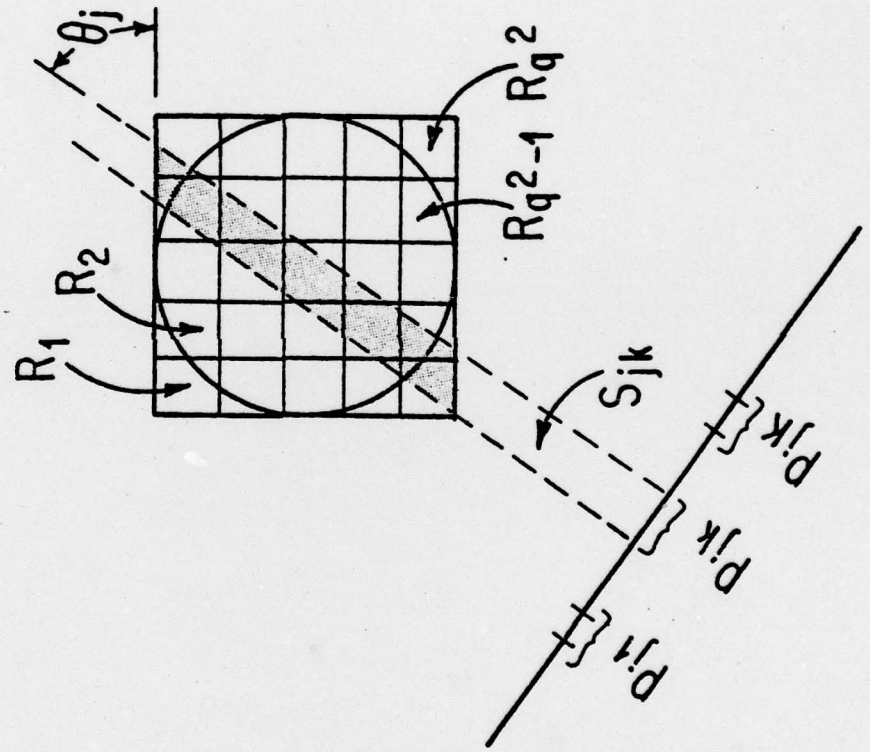


Fig. 5--Spatial discretization of reconstruction domain and line integral data (after Gordon 1974).